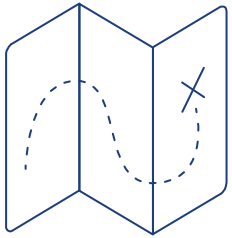
Tiny Tweaks:
Small Interventions to Help Improve Student Worker Performance

Shannon Willis, Digital Projects Lab Manager
University of North Texas



Context

Digital Projects Lab at the University of North Texas

- Digitization lab
- Repositories
  - The Portal to Texas History
  - UNT Digital Library
- Lots of scanners and some cameras

Student Workers in Digital Projects Lab

- Do most of digitizing
- Being the primary creator of images makes them highly valuable
- 10 to 25 imaging student assistants
- 2 librarian supervisors

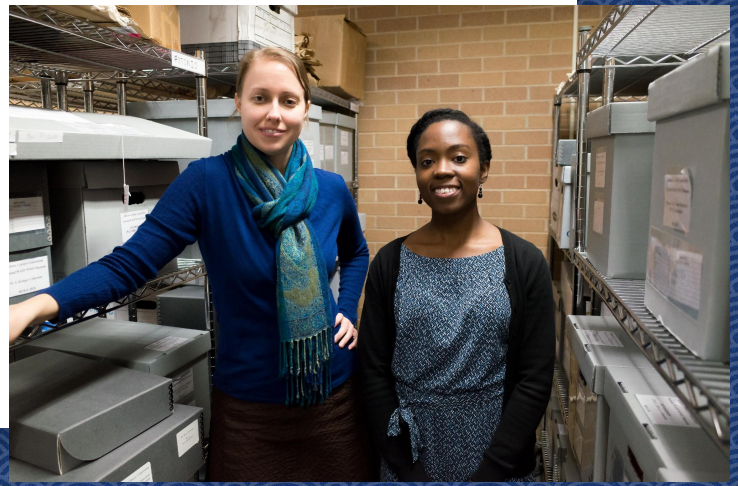


Photo by Christina Kellum, UNT

Student Worker Caveats

- Less work experience
- Little direct imaging experience
- Many other greater priorities
- Part-time (average 19 hours per week)




Vs. What Our Student Assistants Have to Learn and Remember

- 9 different capture systems
- 9 different capture software
- 4 other post-processing software programs
- 9 different categories of material types
- Common problems for each material type and capture system
- File/folder organizational structure
- 4 different file naming conventions
- General computer skills

Small changes, big impact

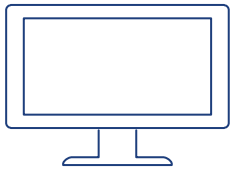
With little time and effort, tiny tweaks can greatly improve student worker performance

Categories

- Giving more information
- Documenting and reporting
- Providing more tools
- Rethinking the workflow

Giving more information

And being super explicit



Beefing up the Project Page

Restoring its glory

A wiki Project Page

1. 042019 - Date:2019-04-09 - Sandborn Books 2018-19
2. 042019 - Date:2019-04-05 - NWTX United Methodist Summerlee Grant
3. 032019 - Date:2019-03-22 - WASP Batch 6
4. 032019 - Date:2019-03-13 - Archives 1 KXAS scrip in box 52
5. 032019 - Date:2019-03-12 - Archives Texas Stonewall Democratic Caucus Papers Box 5
6. 032019 - Date:2019-03-08 - Archives KXAS Box 6
7. 032019 - Date:2019-03-07 - Archives KXAS Box 7
8. 032019 - Date:2019-02-25 - Archives KXAS Box 5, 8
9. 022019 - Date:2019-02-22 - RTH18 Sam Rayburn House
10. 022019 - Date:2019-02-04 - Stewart Title Company 20190204 (Batch4)
11. 012019 - Date:2019-01-18 - RTH18 Lakeway Heritage Center
12. 012019 - Date:2019-01-11 - RTH18 Tarleton State University
13. 012019 - Date:2019-01-04 - Archives Tulia, Brady, and Memphis Collection (AR0915)
14. 012019 - Date:2019-01-03 - Fire Museum of Texas
15. 122018 - Date:2018-12-17 - WASP Batch 5 (201812)
16. 112018 - Date:2018-11-19 - Palacios Book 3
17. 112018 - Date:2018-11-16 - RTH18 Private Collection of Marion Marcel Salas
18. 112018 - Date:2018-11-15 - RTH18 Cathedral Church of St. Matthews
19. 112018 - Date:2018-11-05 - San Antonio PL Funeral Programs (Batch 201811)
20. 112018 - Date:2018-11-01 - Tarrant Count Black Genealogy and Historical Society, Inc.

page | discussion | edit | history

Caseta 18

Contents [hide]

Project Metadata [edit]

Project Info [edit]

Project Title
RTH18 CASETA

Collection Code
Rescuing Texas History, 2018

Institution Code
SG18 - Rescuing Texas History 2018

Institution Code
CASETA - Center for the Advancement and Study of Early Texas Art

Information added

Project Description

Project Description [edit]

- 20 items (3 photos, 2 documents, 15 bulletins)
- 31 capacity points
- "The Emigrant Journey" is a handwritten account of the 1877 experiences of Charles Rudolph, then age 18, and his 15-year-old brother Nelson during their move from southern Illinois to northern Texas, written by Charles in 1878. It is a first-hand, primary source of information on travel to and conditions in north Texas in the late 1800s. The Flora Clarke and Leonard Bennett scrapbooks reflect the experiences of male and female college students in the 1923-1926 era at John Tarleton Agricultural College, then a small, rural two-year Texas school. Contents include photographs of the subjects, their friends and families, basketball teams, football games, military cadets on the drill field, other students, and campus scenes. Clarke's scrapbook also includes notes from friends and male admirers, and memorabilia such as a yell book, newspaper clippings, programs for events, souvenirs from parties and banquets, greeting cards and calling cards, song lyrics, and a poem Clarke wrote about the college. The Edna Wolfe 1925 "Purple Book" is a handbook given to students in that period containing the college's rules and regulations. The college's bulletins (catalogs and announcements) dated from 1921 through 1927 provide context for Bennett's, Clarke's, and Wolfe's experiences, with their photographs and descriptions of the college's history, facilities, faculty, costs, curricula, and courses. These early bulletins also include directory information (names and hometowns) of enrolled students, which is often the only evidence that a summer session student attended the college, as they do not appear in yearbooks.

Project Workflow/Naming Instructions

Project Workflow [edit]

- Naming: (folders have been pre-made in 00.toScan)
  Folder - TSU_ID
  Image - TSU_ID_01 or 02, etc. or magick number

Imaging Workflow and What to watch out for

Imaging Workflow/Instructions [edit]

1. Disbind Bulletin book
2. scan photos on EPSON
3. scan 1877 on Quartz - it is very old so handle with care
4. scan disbound documents and booklet on Copibook

Watch Out For [edit]

Materials that are old and/or fragile **HANDLE WITH CARE!!** If you believe having the object lie flat or having glass pressed on it will damage it, DON'T DO IT. Mark it as too fragile for current workflow/equipment and inform staff that you have concerns about the object.

Copibook [edit]

- Copibook
  - Exposure
    Make sure your image is not over exposed and the color is reasonable. The Copibook calibration profiles go bad quickly and they seem to slowly get brighter and brighter as they die. You should check the exposure with a white matte board to make sure the numbers are good before you start. But you should continue to pay attention to the documents or pages you are scanning as you work to make sure they look reasonable.
  - Glare Continued
    Watch your pages as you scan to make sure you are getting any glare on them. The lighting of the copibook can sometimes hit a bend or wrinkle in the paper and cause a bright highlight on what is otherwise a dull, matte sheet of paper. If it is doing this, lower the exposure.
  - Borders
    Make sure your images have even and reasonably sized borders. You don't need to make them super small, but resist the temptation to make them big. They should also be even on all sides of a page and from page to page within an object. The borders should be correct off the machine; no photoshopping should be needed.
  - Photographic Prints
    Absolutely **NO** "real" photos on the Copibook. A "real" photo, or photographic print, is one that has been printed on photographic paper (as opposed to printed on regular paper like you might see in a book or newspaper). If you are scanning a book, especially one that looks scrapbook-like, check it to make sure that there are not photographic prints in it before beginning. We will likely want the whole object done on a different scanner if "real" photos are included in the object.

page | discussion | edit | history

Common Imaging Errors

This is a list of common imaging errors and mistakes we experience in this lab.

Contents [hide]

Condition [edit]

- Fragility

  Materials are old and/or fragile <<describe the condition issues seen>>. **HANDLE WITH CARE!!** If you believe having the object lie flat or having glass pressed on it will damage it, DON'T DO IT. Mark it as too fragile for current workflow/equipment and inform staff that you have concerns about the object.

- Onion paper, unbound

  Onion paper, and other kinds of super thin, see-through paper, should be scanned on a white background so that the text is clearly visible and the color of the paper does not change dramatically. Scan these pages either on an Epson, or lay down a sheet of white paper on the Copibook, Quartz, or Phase One. The white sheet should be the full background of the image, NOT just the size of the object.

- Onion paper, bound

  When onion paper, and other kinds of super thin, see-through paper, are found in a bound work, a white backer page should be placed behind the paper so that the text is clearly visible and pages behind it don't bleed through. The white backer page should be the same size as the paper and, as best as possible, should not show around the edges. The white back page should only be used when scanning the front side of the page; you do not need to use a white backer page when scanning the back of the page (this helps people see that it is onion skin).

- Glossy photos, especially with white borders:

  When scanning on the Epson, these will often create glares on their white borders. You have to be very aware of which way you are placing the photograph so that the bar of light does not hit a curve in the paper and cause a glare. Glares can either be a bright white or yellow, or they can sometimes even be red. Pay attention when scanning both front and backs of such photos and watch for glares.

- White/light colored paper and white/light color photograph backs:

  When scanning on Epson, make sure that these are not getting over exposed. Some of the Epsons will cut in too much on the whites and you can loose the edges of the paper or photograph. If your Epson is doing this, trying opening the histogram in Epson an pulling the white arrow all the way to the right.

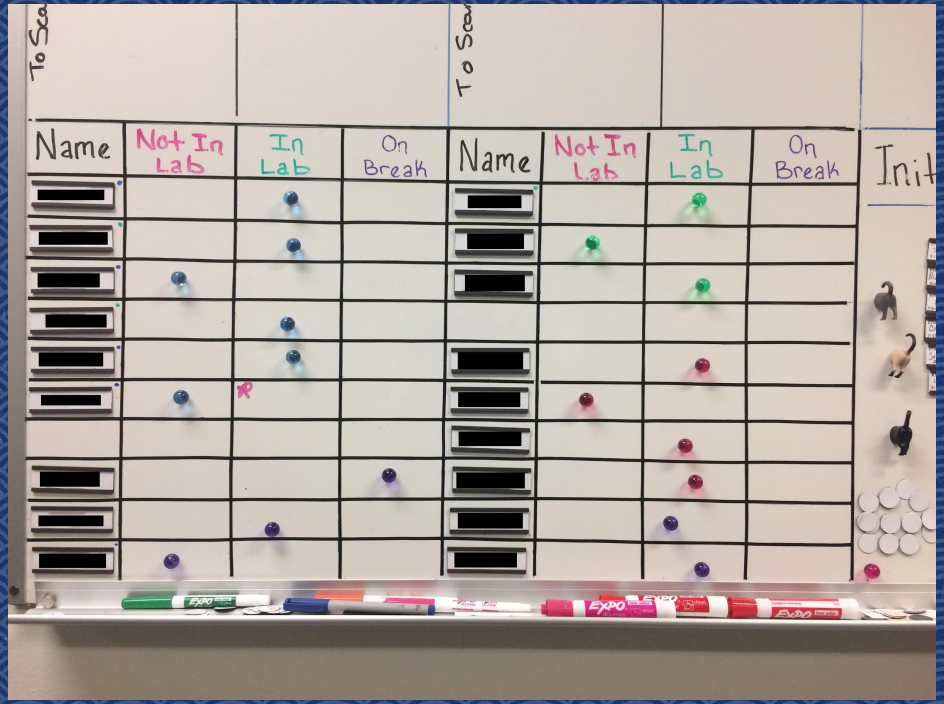
- Halftone and other Printed images (printed as in it clearly came off a printer and was not made on photographic paper)

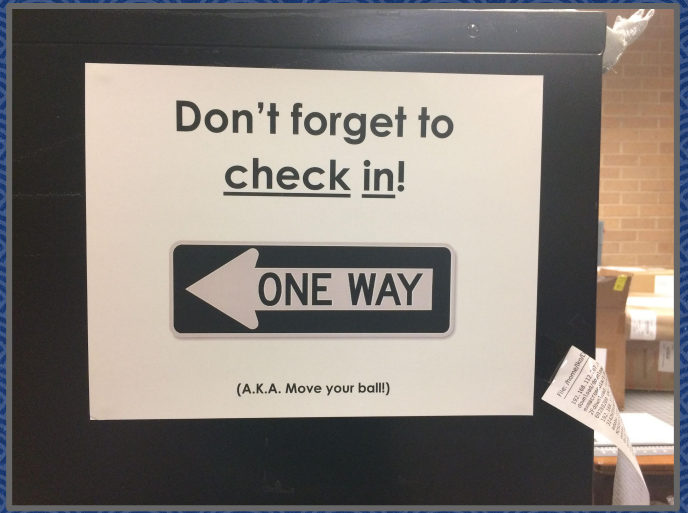
  These often cause moire. Pay particular attention while you are scanning. You may need to raise the resolution you are scanning at or change scanners to avoid the moire pattern (which comes in many shapes and colors but is never desired).

Signage

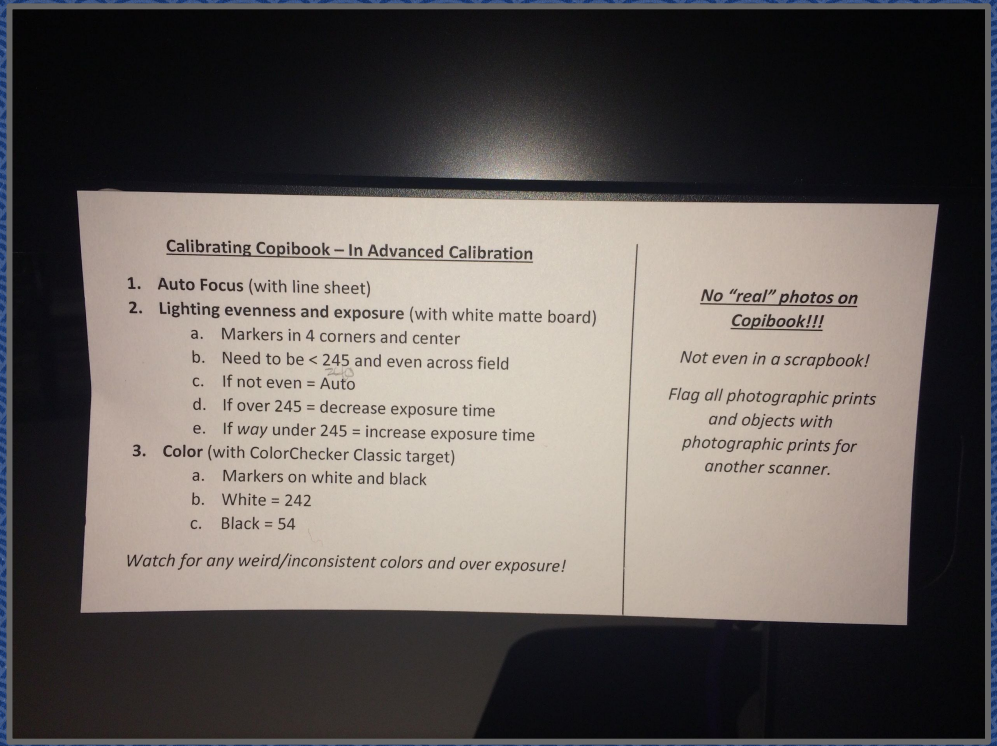
Everywhere

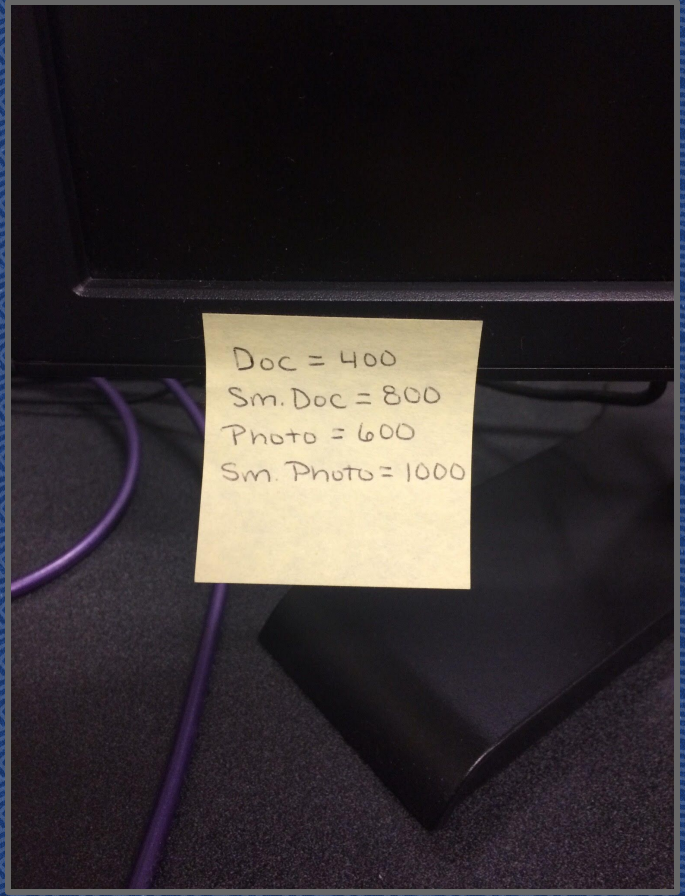
Check-in board and reminder signs



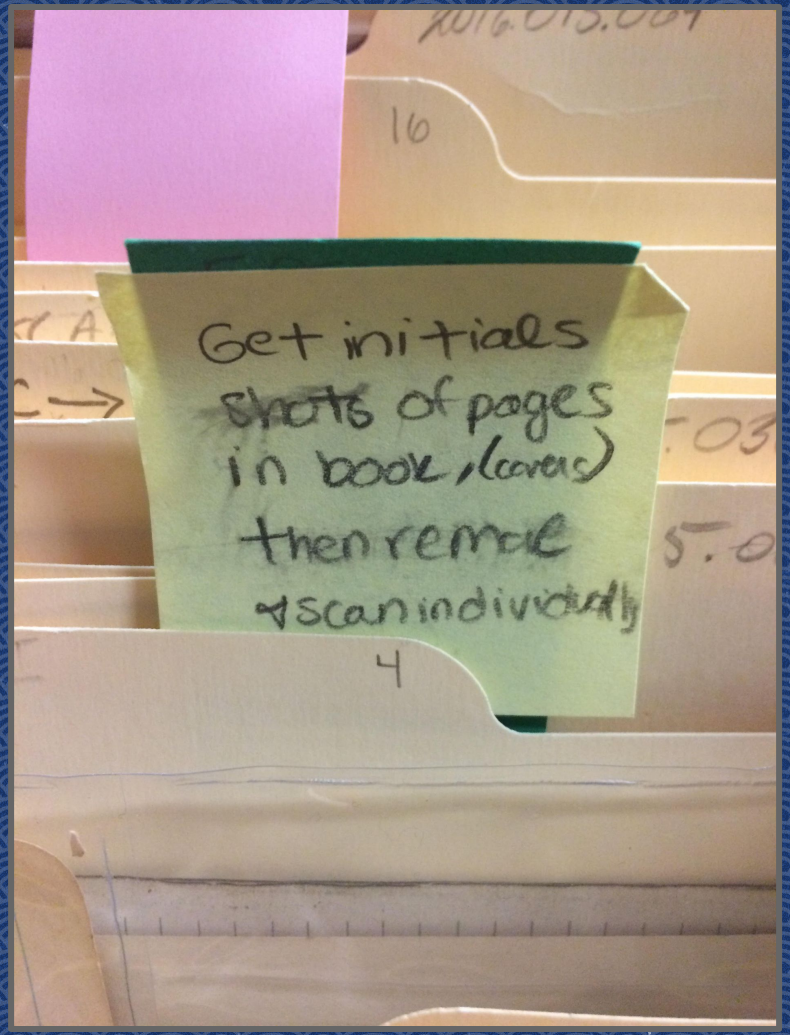


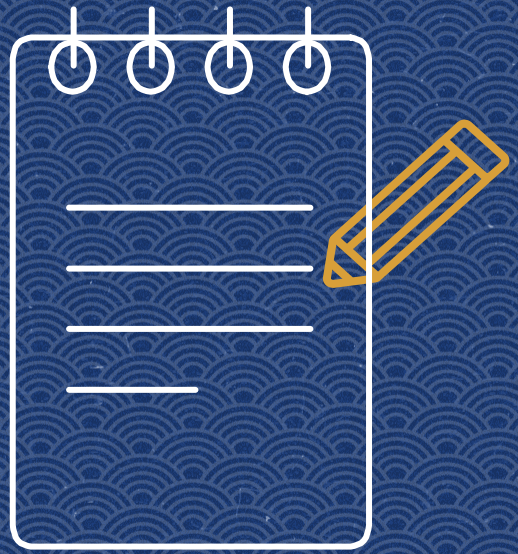
Notes on computers





Notes on physical items





Documenting and reporting

All the things

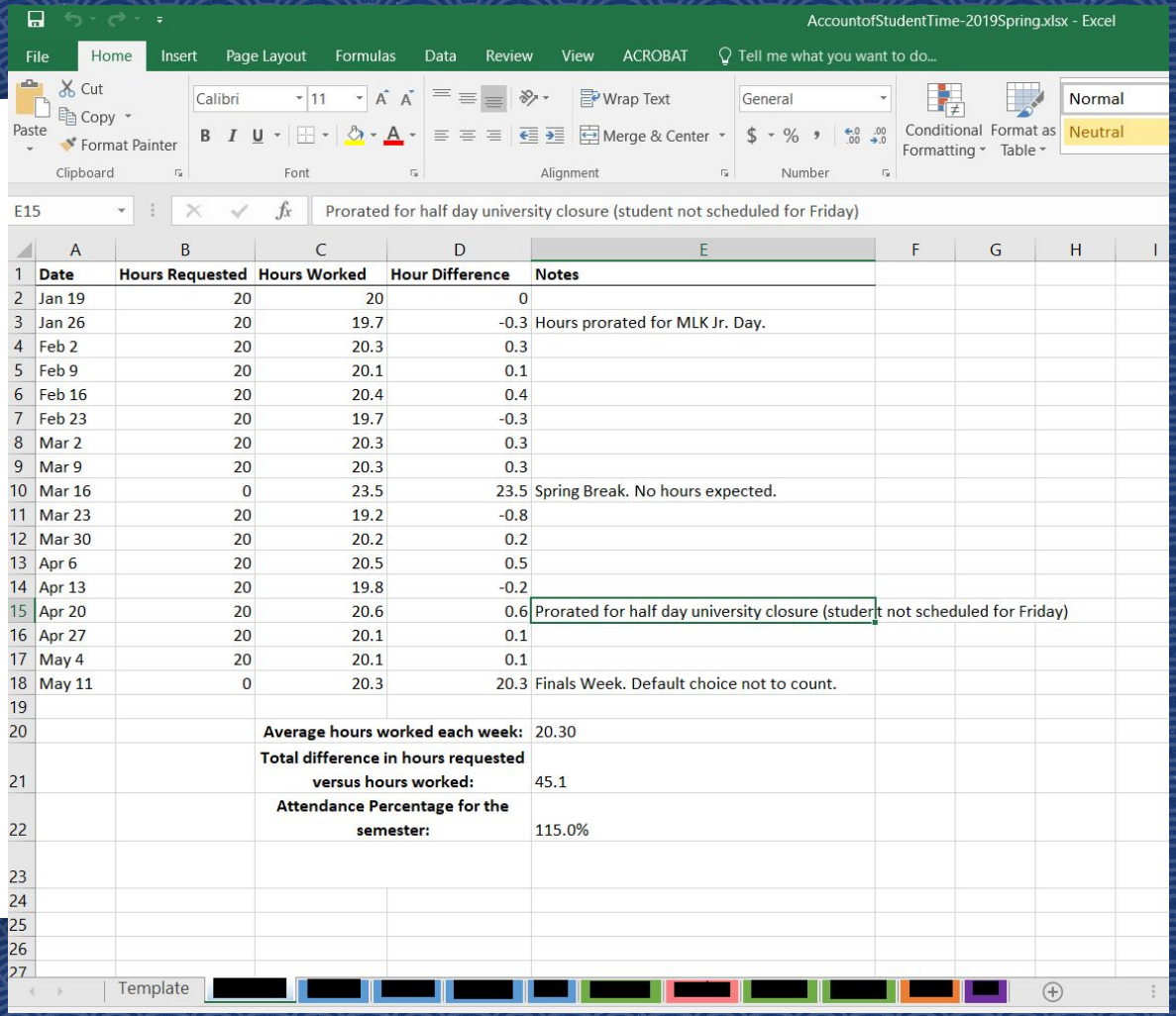
Time Reports

To ensure work attendance

How attendance works in the lab

- Student choose their own schedules
- They enter it in themselves on a spreadsheet each semester
- We expect them to work a minimum of 80% of their scheduled hours
- Less than 80% attendance two semesters in a row results in termination
- More than 95% of their scheduled hours makes them eligible for a raise

Time Tracking



AccountofStudentTime-2019Spring.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View ACROBAT Tell me what you want to do...

Cut Copy Format Painter Paste Clipboard

Calibri 11 Font

Wrap Text Merge & Center Alignment

General Number

Conditional Formatting Format as Table

Normal Neutral

E15 fx Prorated for half day university closure (student not scheduled for Friday)

	A	B	C	D	E
1	**Date**	**Hours Requested**	**Hours Worked**	**Hour Difference**	**Notes**
2	Jan 19	20	20	0	
3	Jan 26	20	19.7	-0.3	Hours prorated for MLK Jr. Day.
4	Feb 2	20	20.3	0.3	
5	Feb 9	20	20.1	0.1	
6	Feb 16	20	20.4	0.4	
7	Feb 23	20	19.7	-0.3	
8	Mar 2	20	20.3	0.3	
9	Mar 9	20	20.3	0.3	
10	Mar 16	0	23.5	23.5	Spring Break. No hours expected.
11	Mar 23	20	19.2	-0.8	
12	Mar 30	20	20.2	0.2	
13	Apr 6	20	20.5	0.5	
14	Apr 13	20	19.8	-0.2	
15	Apr 20	20	20.6	0.6	Prorated for half day university closure (student not scheduled for Friday)
16	Apr 27	20	20.1	0.1	
17	May 4	20	20.1	0.1	
18	May 11	0	20.3	20.3	Finals Week. Default choice not to count.
19					
20			**Average hours worked each week:**		20.30
21			**Total difference in hours requested versus hours worked:**		45.1
22			**Attendance Percentage for the semester:**		115.0%
23					
24					
25					
26					
27					

Template [redacted]

Success of Periodic Emails Throughout Semester

Year before initiative:

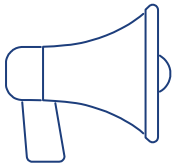
- 8 of 21 worked all their hours.
- 4 of 21 worked at least 80%
- 9 of 21 worked insufficient hours.

Year after initiative:

- 9 of 24 students worked all their hours.
- 12 of 24 worked at least 80%
- 3 of 24 worked insufficient hours.

This Year:

All imaging students worked over 100% of their hours!



Project feedback

Quantitative and qualitative

Quality Control Notes

FinalQC-rth16-chappell.txt - Notepad

File Edit Format View Help

Final QC, RTH16 Chappell Hill -

What I did:

Changed pages to be double-padded (we don't triple pad) or if there were 50+ images, changed them to sequence numbers, or magicknumbered them if they were paginated

Rez'ed down 11 images (3 were at 600, which is an understandable mistake, but the others were at really strange numbers, like 401 and 414. I do not know what happened, but that is strange. Seems to be mostly the rescans too.)

Took out green cast (it bothers me)

Rotated 7 pages

Deleted two duplicate pages

Notes:

In printed items, don't forget we rotate pages to make them readable.

We generally do not triple pad the page extensions. It's either double pad or sequence numbers. The rule of them I have been using is if an item has 50+ images, then I use sequence numbers over the folder-name_01/02 convention.

What needs to be done:

A1997-1-28g: Missing pages 18 and 19
A1997-1-29-1899-1900: I do not know what happened, but 7 pages are under 400 dpi at weird numbers. Rescan.

Student Lead Quality Control Notes

PreQC Note Template [edit]

```
Pre QC Check, Name and Organization
[Project Name]
Start Date:
End Date:
Time:
[PreQC Student Initials]
-------------
# Total Items, # Total TIFs
Total TIFs in Error: # ( % TIF Error Rate)
Total Items in Error: # ( % Item Error Rate)
TIFs sent back for fixes: # ( % TIF Fix Rate)
Items sent back for fixes: # ( % Item Fix Rate)
-------------
ITEMS PREQC STUDENT FIXED:
Example:
Reordered 3 files
Remagicknumbered 1 item (56 files)
-------------
ITEMS TO FIX BY SCANNER:
Example:
009-003: missing page 3. scan.
010-021: missed item. scan.
-------------
NOTES OR COMMENTS:
```

Tracking error and fix rates

TIFs/Items in Error = All that have been marked for fixes as well as all those that have been fixed by QC person.

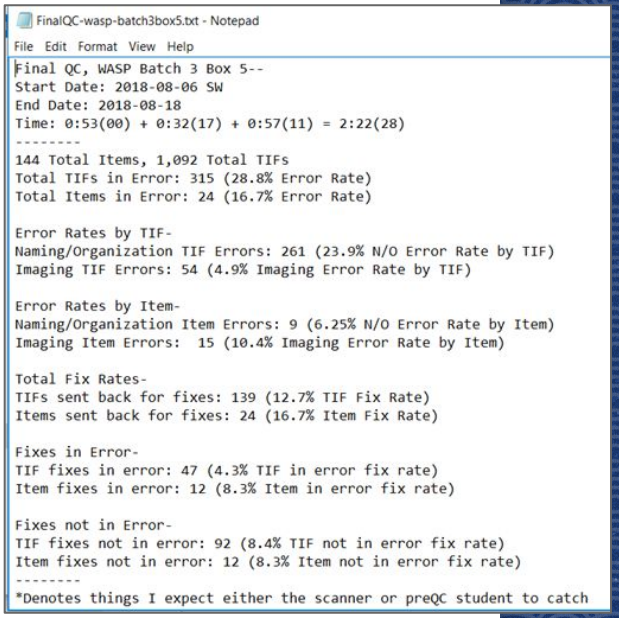
Naming/Organization Error = Files and folders are named incorrectly, organized incorrectly, and/or images are not in the proper order.

Imaging Error = Any error with the actual image file, be that visual or technical

Fix Rate = all images and items sent for fixes

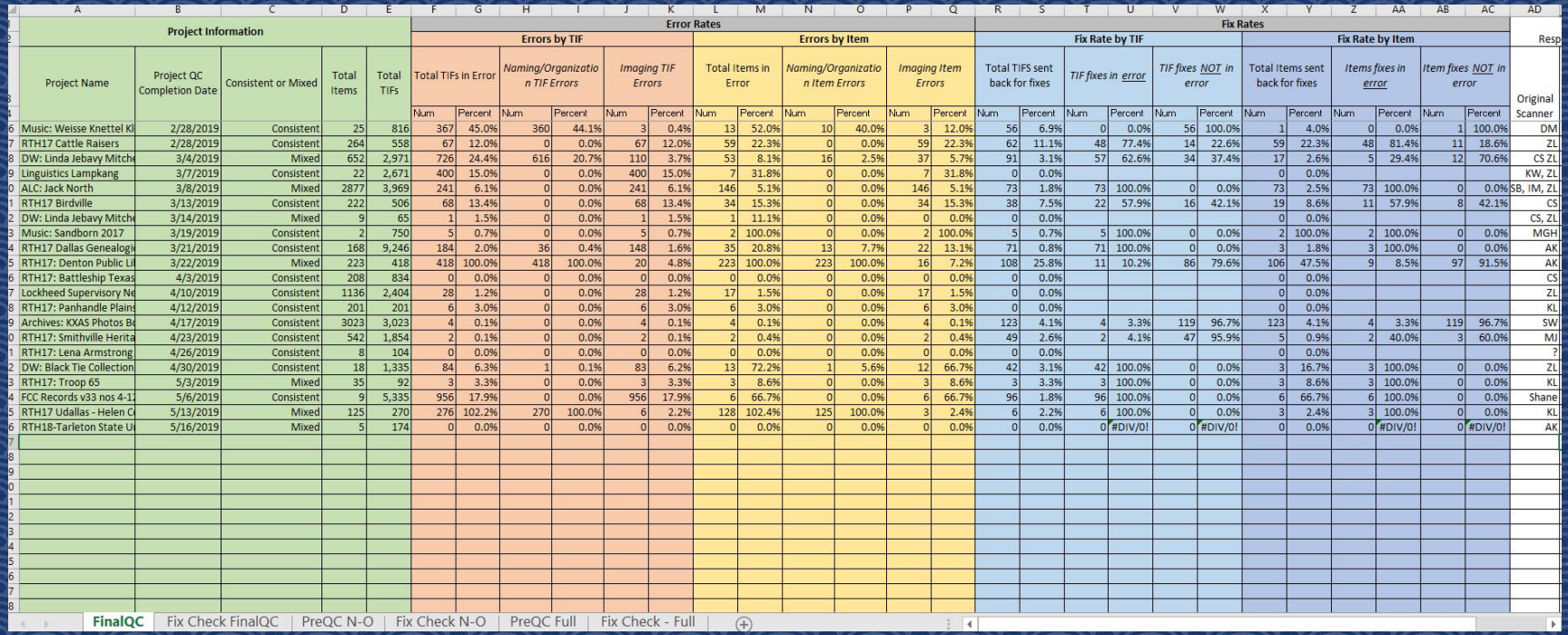
Fixes in Error = all images and items where the fix is something the student should have known about

Fixes not in Error = all images and items requiring fixes that the student cannot have been reasonably expected to know about


FinalQC-wasp-batch3box5.txt - Notepad

File Edit Format View Help

```
Final QC, WASP Batch 3 Box 5--
Start Date: 2018-08-06 SW
End Date: 2018-08-18
Time: 0:53(00) + 0:32(17) + 0:57(11) = 2:22(28)
--------
144 Total Items, 1,092 Total TIFs
Total TIFs in Error: 315 (28.8% Error Rate)
Total Items in Error: 24 (16.7% Error Rate)

Error Rates by TIF-
Naming/Organization TIF Errors: 261 (23.9% N/O Error Rate by TIF)
Imaging TIF Errors: 54 (4.9% Imaging Error Rate by TIF)

Error Rates by Item-
Naming/Organization Item Errors: 9 (6.25% N/O Error Rate by Item)
Imaging Item Errors:  15 (10.4% Imaging Error Rate by Item)

Total Fix Rates-
TIFs sent back for fixes: 139 (12.7% TIF Fix Rate)
Items sent back for fixes: 24 (16.7% Item Fix Rate)

Fixes in Error-
TIF fixes in error: 47 (4.3% TIF in error fix rate)
Item fixes in error: 12 (8.3% Item in error fix rate)

Fixes not in Error-
TIF fixes not in error: 92 (8.4% TIF not in error fix rate)
Item fixes not in error: 12 (8.3% Item not in error fix rate)
--------
*Denotes things I expect either the scanner or preQC student to catch
```



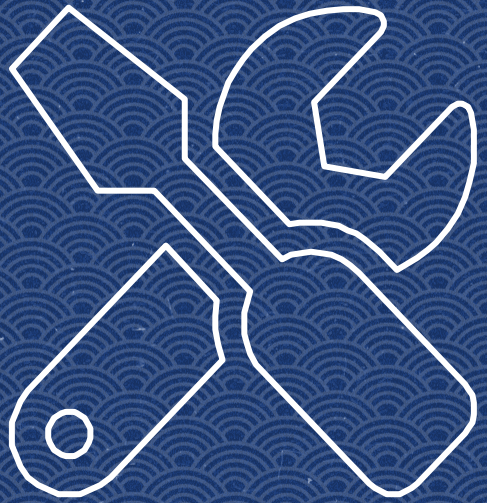
Project Information					Error Rates												Fix Rates												
					Errors by TIF						Errors by Item						Fix Rate by TIF						Fix Rate by Item						Resp
Project Name	Project QC Completion Date	Consistent or Mixed	Total Items	Total TIFs	Total TIFs in Error		*Naming/Organization TIF Errors*		*Imaging TIF Errors*		Total Items in Error		*Naming/Organization Item Errors*		*Imaging Item Errors*		Total TIFS sent back for fixes		*TIF fixes in error*		*TIF fixes NOT in error*		Total Items sent back for fixes		*Items fixes in error*		*Item fixes NOT in error*		Original Scanner
					Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	Num	Percent	
Music: Weisse Knettel Kl	2/28/2019	Consistent	25	816	367	45.0%	360	44.1%	3	0.4%	13	52.0%	10	40.0%	3	12.0%	56	6.9%	0	0.0%	56	100.0%	1	4.0%	0	0.0%	1	100.0%	DM
RTH17 Cattle Raisers	2/28/2019	Consistent	264	558	67	12.0%	0	0.0%	67	12.0%	59	22.3%	0	0.0%	59	22.3%	62	11.1%	48	77.4%	14	22.6%	59	22.3%	48	81.4%	11	18.6%	ZL
DW: Linda Jebavy Mitche	3/4/2019	Mixed	652	2,971	726	24.4%	616	20.7%	110	3.7%	53	8.1%	16	2.5%	37	5.7%	91	3.1%	57	62.6%	34	37.4%	17	2.6%	5	29.4%	12	70.6%	CS ZL
Linguistics Lampkang	3/7/2019	Consistent	22	2,671	400	15.0%	0	0.0%	400	15.0%	7	31.8%	0	0.0%	7	31.8%	0	0.0%					0	0.0%					KW, ZL
ALC: Jack North	3/8/2019	Mixed	2877	3,969	241	6.1%	0	0.0%	241	6.1%	146	5.1%	0	0.0%	146	5.1%	73	1.8%	73	100.0%	0	0.0%	73	2.5%	73	100.0%	0	0.0%	SB, IM, ZL
RTH17 Birdville	3/13/2019	Consistent	222	506	68	13.4%	0	0.0%	68	13.4%	34	15.3%	0	0.0%	34	15.3%	38	7.5%	22	57.9%	16	42.1%	19	8.6%	11	57.9%	8	42.1%	CS
DW: Linda Jebavy Mitche	3/14/2019	Mixed	9	65	1	1.5%	0	0.0%	1	1.5%	1	11.1%	0	0.0%	0	0.0%	0	0.0%					0	0.0%					CS, ZL
Music: Sandborn 2017	3/19/2019	Consistent	2	750	5	0.7%	0	0.0%	5	0.7%	2	100.0%	0	0.0%	2	100.0%	5	0.7%	5	100.0%	0	0.0%	2	100.0%	2	100.0%	0	0.0%	MGH
RTH17 Dallas Genealogi	3/21/2019	Consistent	168	9,246	184	2.0%	36	0.4%	148	1.6%	35	20.8%	13	7.7%	22	13.1%	71	0.8%	71	100.0%	0	0.0%	3	1.8%	3	100.0%	0	0.0%	AK
RTH17: Denton Public Lil	3/22/2019	Mixed	223	418	418	100.0%	418	100.0%	20	4.8%	223	100.0%	223	100.0%	16	7.2%	108	25.8%	11	10.2%	86	79.6%	106	47.5%	9	8.5%	97	91.5%	AK
RTH17: Battleship Texas	4/3/2019	Consistent	208	834	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%					0	0.0%					CS
Lockheed Supervisory Ne	4/10/2019	Consistent	1136	2,404	28	1.2%	0	0.0%	28	1.2%	17	1.5%	0	0.0%	17	1.5%	0	0.0%					0	0.0%					ZL
RTH17: Panhandle Plains	4/12/2019	Consistent	201	201	6	3.0%	0	0.0%	6	3.0%	6	3.0%	0	0.0%	6	3.0%	0	0.0%					0	0.0%					KL
Archives: KXAS Photos Be	4/17/2019	Consistent	3023	3,023	4	0.1%	0	0.0%	4	0.1%	4	0.1%	0	0.0%	4	0.1%	123	4.1%	4	3.3%	119	96.7%	123	4.1%	4	3.3%	119	96.7%	SW
RTH17: Smithville Herita	4/23/2019	Consistent	542	1,854	2	0.1%	0	0.0%	2	0.1%	2	0.4%	0	0.0%	2	0.4%	49	2.6%	2	4.1%	47	95.9%	5	0.9%	2	40.0%	3	60.0%	MJ
RTH17: Lena Armstrong	4/26/2019	Consistent	8	104	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%					0	0.0%					?
DW: Black Tie Collection	4/30/2019	Consistent	18	1,335	84	6.3%	1	0.1%	83	6.2%	13	72.2%	1	5.6%	12	66.7%	42	3.1%	42	100.0%	0	0.0%	3	16.7%	3	100.0%	0	0.0%	ZL
RTH17: Troop 65	5/3/2019	Mixed	35	92	3	3.3%	0	0.0%	3	3.3%	3	8.6%	0	0.0%	3	8.6%	3	3.3%	3	100.0%	0	0.0%	3	8.6%	3	100.0%	0	0.0%	KL
FCC Records v33 nos 4-12	5/6/2019	Consistent	9	5,335	956	17.9%	0	0.0%	956	17.9%	6	66.7%	0	0.0%	6	66.7%	96	1.8%	96	100.0%	0	0.0%	6	66.7%	6	100.0%	0	0.0%	Shane
RTH17 Udallas - Helen C	5/13/2019	Mixed	125	270	276	102.2%	270	100.0%	6	2.2%	128	102.4%	125	100.0%	3	2.4%	6	2.2%	6	100.0%	0	0.0%	3	2.4%	3	100.0%	0	0.0%	KL
RTH18-Tarleton State U	5/16/2019	Mixed	5	174	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	#DIV/0!	0	#DIV/0!	0	0.0%	0	#DIV/0!	0	#DIV/0!	AK

FinalQC | Fix Check FinalQC | PreQC N-O | Fix Check N-O | PreQC Full | Fix Check - Full

Error Stats Spreadsheet

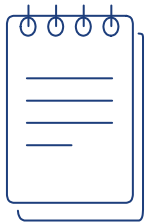
Benefits of Statistics

- Gives feedback more power
- Supports critical feedback
- Emphasizes positive feedback
- Supervisors can see where *they* can improve
- Numbers are sombering
- Improves training and tutorials



Providing more tools

For success



Checklists

- With training
- With tasks
- With equipment

With tasks

Quality control checks by Student Leads

page | discussion | edit | history

HowToOther PreQCImages

Contents [hide]

Overview of Workflow

1. Check folder structure and folder names
2. Check file names
3. Check bit depths
4. Check resolutions
5. Check color profiles
6. Check that every item has been scanned and visual checks

Project Name:_____ Initials_____

Name and Organization PreQC Checklist: Bound Collections

- ☐ Each item has its own folder
- ☐ Item folders are not nested (unless specified for the project)
- ☐ For scantailored items, "Originals" subfolders are present
- ☐ Item folder names match the inventory and/or identifier on item
- ☐ There are no typos or other errors in the folder names
- ☐ Each individual item has been scanned
- ☐ All pages of each individual item have been scanned
- ☐ For scantailored items, no text on the page has been cut off
- ☐ All scanned pages are in the correct order (i.e., match the order of the physical item)
- ☐ Sequence numbers, magicknumbers, or special numbering have been chosen correctly
- ☐ The magicknumbering and/or sequence numbers are correct when used
- ☐ Any special numbering file is accurate and a .txt file

Project Name:_____ Initials_____

Name and Organization PreQC Checklist: Unbound/Mixed Collections

- ☐ Each item has its own folder
- ☐ Item folders are not nested (unless specified for the project)
- ☐ Item folder names match the inventory and/or identifier on item
- ☐ The Partner Code Prefix is included in each item name
- ☐ There are no typos or other errors in the folder names
- ☐ Filenames match the folder names (expect when magicknumbering or sequence numbers are appropriate)
- ☐ The correct naming convention has been chosen for each item
- ☐ Items that require magicknumbering or sequence numbers have been named correctly and accurately
- ☐ Each individual item has been scanned
- ☐ All sides, parts, and pages of each individual item have been scanned
- ☐ All images are in the correct order (i.e., match the order of the physical item and they follow our standards for Image Sequence)

With equipment

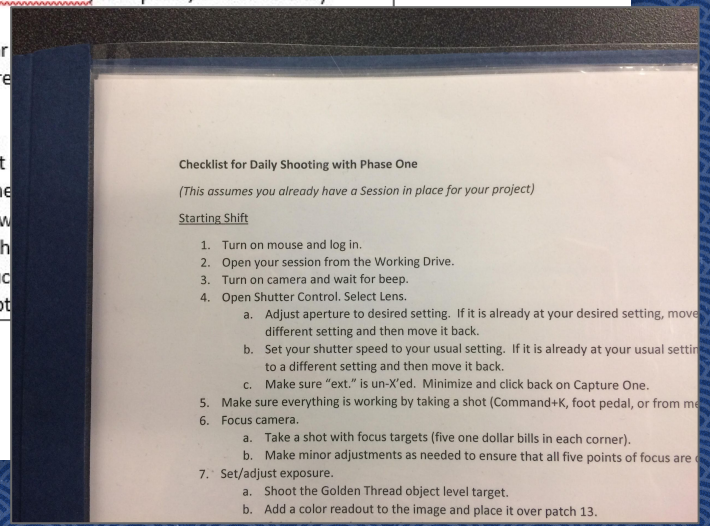
Phase One Checklists

Checklist for Daily Shooting with Phase One

(This assumes you already have a Session in place for your project)

Starting Shift

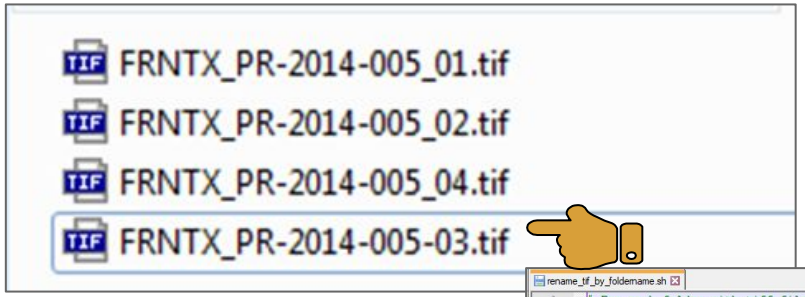
1. Turn on mouse and log in.
2. Open your session from the Working Drive.
3. Turn on camera and wait for beep.
4. Open Shutter Control. Select Lens.
    a. Adjust aperture to desired setting. If it is already at your desired setting, move it to a different setting and then move it back.
    b. Set your shutter speed to your usual setting. If it is already at your usual setting, move it to a different setting and then move it back.
    c. Make sure "ext." is un-X'ed. Minimize and click back on Capture One.
5. Make sure everything is working by taking a shot (Command+K, foot pedal, or from menu).
6. Focus camera.
    a. Take a shot with focus targets (five one dollar
    b. Make minor adjustments as needed to ensure
7. Set/adjust exposure.
    a. Shoot the Golden Thread object level target.
    b. Add a color readout to the image and place it
    c. If the color readout number is not close to the
        i. FIRST adjust shutter speed (up or dow
        ii. SECOND, make fine adjustments to th
        the numbers even closer (avoid reduc
        increasing exposure above 0.2 in Capt



Command line tools

Also “Rethinking the workflow”

A solution to file naming errors



rename_tif_by_foldername.sh

```
# For each folder with tiff files
for f in * ; do
  # move into the folder
  cd "$f"

  # Start a counter variable called a
  a=1

  # Loop through all .tif files in the folder
  for i in *.tif; do

    # prepare the format we want to use, right now it is the same as the folder name $f
    new=$(printf "${f}_%02d.tif" "$a") #02 pad to length of 2

    # move the file to the new name
    mv -i -- "$i" "$new"

    # increment the counter
    let a=a+1

  # close this inner loop.
  done

  # Show that we are finished with this folder
  echo "$f renamed"
  # Move out of the folder going one level back.
  cd ..

# finish the script
done
```

Rename Files [edit]

Digital Libraries has written a Bash program to rename all files according to their folder name. To use the program, all you have to do is call it. Using this program will allow you to rename files by their folder names (the base structure naming convention) in large batches.
To do so:

- From the command line, navigate to the directory (folder) with folders that have files you need to rename (this should be either the Scanning or Processing folder in your project)
  - Open Cygwin terminal
  - Navigate to the P drive

```
cd P:
```

cd stands for change directory

- Navigate to the appropriate workflow folder in your project folder on the P: drive. For example:

```
cd DigiProj/scanned_for_kempner/2.toProcess
```

Note: The code is expecting to find a bunch of folders in the location it is in. If you are within an item folder and it just finds files, the code will not work. This code is meant for renaming files in a batch of item folders, not for renaming files in a single item folder.

- Call the code by running this:

```
bash ../../Software/imaging-scripts/rename_tif_by_foldername.sh
```

Note: You may have to adjust your path (the ../../ part of the code) depending on how nested your project is. Each ../ part tells the computer to go a step up the folder structure. Then when you give it a name of a directory (like "Software") it starts to go down in the folder structure. So, for example, if you need it to go up more than two steps, you will need another ../ in addition to the two written above (../../../Software/imaging-scripts/rename_tif_by_foldername.sh). If you get the path wrong, you won't break anything. You will just get an error message that the file you are trying to call is not found.

- After running the code, check the items you renamed to make sure you do not have any items that should have sequence numbers or magicknumbers instead of the base structure

page | discussion | edit | history

QC At The Command-Line

Contents [hide]

2. Check filenames [edit]

At this step, you want to make sure all filenames match folder names for in the project.

For projects of items that will each have only a few images per item, we use the command line to check the filenames. When you check unbound or mixed collections, remember the rule that if an item has over 50 images, it should use sequence numbers and NOT the identifier base structure.

Procedure:

- Open Cygwin Terminal
- Navigate to the P drive

```
cd P:
```

cd stands for "change directory"

- Naviagate to the PreQC folder inside the project folder you are preQC-ing

```
cd DigiProj/scanned_for_kempner/3.toPreQC
```

- Use the following code to check if the filenames match the folder names

```
for i in *; do echo -n $i ":" ; ls "$i" --ignore="Thumbs.db" | grep -v -E "^$i\_[0-9]{2}\.tif$" | wc -l ; done | grep -v ":0"$
```

- Also use the following code to check for empty folders

```
for i in *; do echo -n $i ":" ; ls "$i"/*.tif | wc -l ; done | grep ":0"
```

Rethinking the workflow

Again and again

Pre-making item folders

Before a student starts work on a project

Pre-made folders

Make a Batch of Folders [edit]

This command lets you make of batch of folders that will all have the same stem to their name

Programs Used:

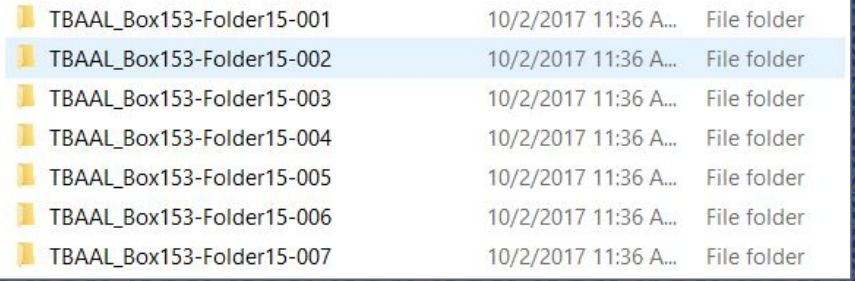
- mkdir

```
for i in {[Start number]..[Stop number]} ; do mkdir [Stem of Name]-$i; done
```

Example:

```
for i in {001..045} ; do mkdir TBAAL_Box153-Folder15-$i; done
```

Example Result:



Make a Batch of Folders from a Text File of Names [edit]

This command lets you make of batch of folders using a text file that has the names listed on different lines

Programs Used:

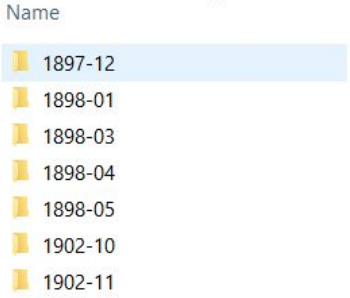
- read
- mkdir

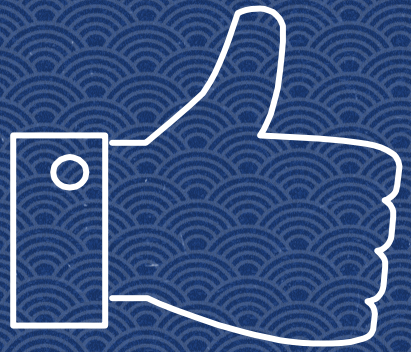
```
<[Name of your file] tr -d '\r'| while read -r l || [ -n "$l" ] ; do mkdir $l ; done
```

Example:

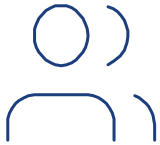
```
<ItemNames.txt tr -d '\r'| while read -r l || [ -n "$l" ] ; do mkdir $l ; done
```

Example Result:





- Documenting and reporting
- Rethinking the workflow
- Giving more information
- Providing more tools



Not an exhaustive list

The job is never done

Thank you!

Questions?

Shannon.Willis@unt.edu

Presentation theme from SlidesCarnival